

Docket No.: CL000685-CON
Serial No.: TO BE ASSIGNED
Inventors: YAN, Chunhua et al.
Title: ISOLATED HUMAN DRUG-METABOLIZING...

```
1 TTTCTTCTGT TTGCTTACTC CCTATCCGGG GGGCCCAAGGC GCTGTCTCCG
51 CCGCCCAAGC CCGCGTAAAC CCTGGGTGAC CTCGGAGACA TCCGTTGGAG
101 CATGAGTTCC CGACATCAGG CGGCGGCGGT GGTCCGGGAG AAACCCGGCG
151 GCGGGGAGAT AAGCCTGCCC AGGAGGCAGG GGGCTGGGCT AGCTGCCCCG
201 CCGCGCGCCT GACTTCGTTG GGGAGGGAGA CGCCCGGCTC CCGCCCTAA
251 CTAGCCACG CGCGCGGAGC GCCTGGGAGA GGAGAAGGAG CCGACCTGCC
301 GAGATGGAGG CGACCGGCAC CTGGGCGCTG CTGCTGGCGC TGGCGCTGCT
351 CCTGCTGCTG ACGCTGGCGC TGTCCGGGAC CAGGGCCCGA GGCCACCTGC
401 CCCCCGGGCC CACGCCGCTA CCACTGCTGG GAAACCTCCT GCAGCTACGG
451 CCGGGGGGCG TGTATTCAGG GCTCATGCGG CTGAGTAAGA AGTACGGACC
501 GGTGTTTACC ATCTACCTGG GACCTGGCG GCCTGTGGTG GTCCTGGTTG
551 GGCAGGAGGC GTGCGGGAG GCGCTGGGAG GTCAGGCTGA GGAGTTCAGC
601 GGCCGGGGAA CCGTAGCGAT GCTGGAAGGG ACTTTTGATG GCCATGGGGT
651 TTTCTTCTCC AACGGGGAGC GGTGGAGGCA GGTGAGGAAG TTTACCATGC
701 TTGCTCTGCG GGACCTGGGC ATGGGAAGC GAGAAGGCGA GGAGCTGATC
751 CAGGCGGAGG CCGGTGTCT GGTGGAGACA TTCCAGGGGA CAGAAGGACG
801 CCCATTTCAT CCTTCCCTGC TGCTGGCCCA GGCCACTCC AACGTAGTCT
851 GTCCTCTCCT CTTTGGCCTC CGCTTCTCCT ATGAGGATAA GGAGTTCAG
901 GCGGTGGTCC GGGCAGCTGG TGGTACCCTG CTGGGAGTCA GCTCCAGGG
951 GGGTCAGACC TACGAGATGT TCTCTGGTT CCTGCGGCC CTGCCAGGCC
1001 CCCACAAGCA GCTCCTCCAC CACGTCAGCA CTTTGGCTGC CTTACAGTCC
1051 CGGCAGGTGC AGCAGACCA GGGGAACCTG GATGCTTCGG GCCCGCAGC
1101 TGACCTTGTG GATGCTTCC TGCTGAAGAT GGCACAGGAG GAACAAAACC
1151 CAGGCACAGA ATTCACCAAC AAGAACATGC TGATGACAGT CATTTATTG
1201 CTGTTTGCTG GGACGATGAC GGTGAGCACC ACGGTGCGCT ATACCCTCCT
1251 GTCCTGATG AAATACCCTC ATGTCCAAAA GTGGGTACGT GAGGAGCTGA
1301 ATCGGGAGCT GGGGGCTGGC CAGGCACCAA GCCTAGGGGA CCGTACCCGC
1351 CTCCCTTACA CCGACGCGGT TCTGCATGAG GCGCAGCGGC TGCTGGCGCT
1401 GGTGCCCATG GGAATACCCC GCACCCTCAT GCGGACCACC CGCTTCCGAG
1451 GGTACACCTG GCCCAGGGC ACGGAGGTCT TCCCCCTCTC TGGCTCCATC
1501 CTGCATGACC CCAACATCTT CAAGCACCCA GAAGAGTTCA ACCCAGACCG
1551 TTTCTGGAT GCAGATGGAC GGTTCAGGAA GCATGAGGCG TTCCTGCCCT
1601 TCTCCTTAGG GAAGCGTGTG TGCTTGGAG AGGGCCTGGC AAAAGCGGAG
1651 CTCTTCTCTT TCTTACCAC CATCTACAA GCCTTCTCCC TGGAGAGCCC
1701 GTGCCCGCG GACACCCTGA GCCTCAAGCC CACCGTCAGT GGCCTTTTCA
1751 ACATTCCCCC AGCCTTCCAG CTGCAAGTCC GTCCCACTGA CCTTCACTCC
1801 ACCACGCAGA CCAGATGAAG GAAGGCAACT TGGAAAGTGT GGGTGCCAG
1851 GACGGTGCCT CCAGCCTCAA CAGTGGGCAT GGACAGGGTT AATGTCTCCA
1901 GAGTGATAC TGCAGGCAGC CACATTTACA CGCCTGCAGT TGTTTTCCGG
1951 AGTCTGTCCC ACGGCCACAC CGCTCACTTG ACTCATGCTG CTAAGATGCA
2001 CAACCGCACA CCCATACACA ACTACAAGGG CCACAAAGCA ACTGCTGGGT
2051 TAGCTTTCCA CAGACATAAA TATAGTCCAT CTGCAATCAC AAGCACATAG
2101 CCAGGTAACC CACCAACTCC CCTGGATCTG CAGCCACAC GTGGGAGTCT
2151 GGCTGTACCT TTCACAAGCC ACAGAAACGG CCACACATGT TCACAGTCA
2201 CACGCCCTCT CCATTATCG AACTTCTCAG TGTCCCTGTC CCTGGTGCCT
2251 GGCACAGGGA ACAGCATGCC CCTCCGGGG TCATGCCACC CAGAGACTGT
2301 CGTGTCTAT GGCCCAACT CATGCTCCCT CTCTTGGCTA CACCACTCTC
2351 CCAGCTGTG ACCACCATG TCCACACACC CCAACCACT TGTCACACA
2401 GCTACCCACG TACGACATCG TCCTGGCTCC CCAGAGTATC TTCCCACTGA
2451 GACACGCCCG CCCACAGAG GCACAGTCCC CAGCCACCTC TGCAACTGCA
2501 GCCCTCAGTC ACCCTTTTTT AAGCACCTG ATTCTACCA ATGCAACAC
2551 ATCTGGGTCT GCGATTATGC ACAGAGACTT TGGACATACG AGGACCCTCA
2601 GACCGGAGGA ACACCTGCCC AACCCCAACA CGTGCTTATG TAACCACGTG
2651 GAAAGCGGCC CTGCTGCCC CTCCACACAC ACATACACAC TCACTGATCT
2701 ACAGCCCCTG TTCGGCGTCA GAGTCCCCAC TAGACCCAGT GGAAGGGGTT
2751 AGAGACCAAG TAGGGGCCAG TTTCCAATT ACCCTGTGAG GGAGTGAGCC
2801 GGATCTGACG TTCCTTGTGA CTTAAGGGTC CGGCTTGGGA ATTAAGTTT
2851 GTTCTGGCC TTTAGCCTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
2901 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA (SEQ ID NO:1)
```

FEATURES:
5'UTR: 1-303
Start Codon: 304
Stop Codon: 1815
3'UTR: 1818

FIGURE 1A

Docket No.: CL000685-CON
 Serial No.: TO BE ASSIGNED
 Inventors: YAN, Chunhua et al.
 Title: ISOLATED HUMAN DRUG-METABOLIZING...

HOMOLOGOUS PROTEINS:

gi 117254 sp P24461 CPG1_RABIT CYTOCHROME P450 2G1 (CYPIIG1) (P...	516	e-145
gi 404777 gb AAA31432.1 (L10912) cytochrome P-450 2B-Bx [Oryct...	513	e-144
gi 479930 pir S35666 cytochrome P450 2B4 isoform Bx - rabbit >...	513	e-144
gi 117212 sp P00178 CPB4_RABIT CYTOCHROME P450 2B4 (CYPIIB4) (P...	508	e-143
gi 320075 pir S31277 cytochrome P450 2B4-B1 - rabbit >gi 21369...	508	e-143
gi 4731350 gb AAD28466.1 (AF128849) cytochrome P450 2B10 relat...	508	e-143
gi 3123191 sp P04167 CPB2_RAT CYTOCHROME P450 2B2 (CYPIIB2) (P4...	508	e-143
gi 117213 sp P12789 CPB5_RABIT CYTOCHROME P450 2B5 (CYPIIB5) (P...	507	e-142
gi 89973 pir A27717 cytochrome P450 2B5, hepatic (form HP1) - ...	507	e-142
gi 2144292 pir O4RTP2 cytochrome P450 2B2 - rat	506	e-142

Est:

gb BE148597 BE148597 MR0-HT0241-150500-010-b02 HT0241 Homo sapi...	1091	0.0
gb BF359243 BF359243 RC6-ET0072-150600-011-F01 ET0072 Homo sapi...	779	0.0
gb AW753778 AW753778 RC1-CT0286-301099-011-f08 CT0286 Homo sapi...	652	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

gb BE148597	head_neck
gb BF359243	lung_tumor
gb AW753778	colon

Tissue expression:

Human leukocyte

FIGURE 1B

Docket No.: CL000685-CON
Serial No.: TO BE ASSIGNED
Inventors: YAN, Chunhua et al.
Title: ISOLATED HUMAN DRUG-METABOLIZING...

1 MEATGTWALL LALALLLLT LALSGTRARG HLPPTPLP LLGNLLQLRP
51 GALYSGLMRL SKKYGPVFTI YLGPWRPVV LVGQEAVERA LGGQAEFFSG
101 RGTVANLEGT FDHGVFFSN GERWRQLRKF TMLALRDLGM GKREGEELIQ
151 AEARCLVETF QGTEGRPFDP SLLLAQATSN VVCSLLFGLR FSYEDKEFQA
201 VVRAAGGTLL GVSSQGGQTY EMFSWFLRPL PGPHKQLLHH VSTLAAFTVR
251 QVQHQGNLD ASGPARDLVD AFLKMAQEE QNPGTEFTNK NMLMTVIYLL
301 FAGTMTVSTT VGYTLLLLMK YPHVQKWVRE ELNRELGAGQ APSLGDRTLL
351 PYTDAVLHEA QRLLALVPMG IPRTLMTTR FRGYTLPGQT EVFPLLGSIL
401 HDPNIFKHPE EFNPDRLDA DGRFRKHEAF LPFSLGKRVK LGEGLAKAEL
451 FLFFTTILQA FSLESPCPD TLSLKPTVSG LFNIPPAFQL QVRPTDLHST
501 TQTR (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

Prosite search results:

[1] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
CAMP- and cGMP-dependent protein kinase phosphorylation site

128-131 RKFT

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 6

1 61-63 SKK
2 99-101 SGR
3 248-250 TVR
4 288-290 TNK
5 378-380 TTR
6 473-475 SLK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 3

1 119-122 SNGE
2 192-195 SYED
3 343-346 SLGD

[4] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 10

1 51-56 GALYSG
2 109-114 GTFDGH
3 115-120 GVFFSN
4 188-193 GLRFSY
5 207-212 GTLLGV
6 257-262 GNLDAS
7 284-289 GTEFTN
8 339-344 GQAPSL
9 370-375 GIPRTL
10 444-449 GLAKAE

[5] PDOC00009 PS00009 AMIDATION
Amidation site

Number of matches: 2

1 140-143 MGKR
2 435-438 LGKR

[6] PDOC00029 PS00029 LEUCINE_ZIPPER
Leucine zipper pattern

Number of matches: 2

1 32-53 LPPGPTPLPLLGNLLQLRPGAL
2 39-60 LPLLGNLLQLRPGALYSGLMRL

FIGURE 2A

Docket No.: CL000685-CON
Serial No.: TO BE ASSIGNED
Inventors: YAN, Chunhua et al.
Title: ISOLATED HUMAN DRUG-METABOLIZING...

[7] PDOC00081 PS00086 CYTOCHROME_P450
Cytochrome P450 cysteine heme-iron ligand signature

433-442 FSLGKRVCLG

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	5	25	2.305	Certain
2	64	84	0.783	Putative
3	170	190	1.041	Certain
4	292	312	2.031	Certain
5	448	468	1.133	Certain

FIGURE 2B

Docket No.: CL000685-CON
Serial No.: TO BE ASSIGNED
Inventors: YAN, Chunhua et al.
Title: ISOLATED HUMAN DRUG-METABOLIZING...

BLAST Alignment t T p Hit:

>gi|117254|sp|P24461|CPG1_RABIT CYTOCHROME P450 2G1 (CYP11G1)
(P450-NMB) (OLFACTIVE)
pir||B31944 cytochrome P450 2G1 - rabbit
Length = 494

Score = 516 bits (1315), Expect = e-145
Identities = 248/491 (50%), Positives = 345/491 (69%), Gaps = 3/491 (0%)

Query: 1 MEATGTWALLLALAL-LLLLTLALSGTRARGHLPPGPTPLLLGNLLQLRPGALYSGLMR 59
ME G + + LAL LL+ +A + G LPPGPTP+P LGNLLQ+R A + ++
Sbjct: 1 MELGGAFTIFLALCFSCLLILIAWKRVQKPGRLPPGPTPIPLGNLLQVRTDATFQSFLK 60

Query: 60 LSKKYGPVFTIYLGWPVPVVLVQGEAVREALGGQAEFFSGRGTVMLEGTFDGHGVFFS 119
L +KYGPVFT+Y+GP RPVV+L G EAV+EAL +A+EFSGRG +A +E F GHGV +
Sbjct: 61 LREKYGPVFTVYMG-P-RPVVILCGHEAVKEALVDRADFEFSGRGELASVERNFGHGVALA 119

Query: 120 NGERWRQLRKFTMLALRDLGMGKREGEELIQAEARCLVETFGTEGRPFDPSSLQAQTS 179
NGERWR LR+F++ LRD GMGKR EE IQ EA L+E F+ T+G P DP+ L++ S
Sbjct: 120 NGERWRILRRFSLTILRDFGMGKRSIEERIQEEAGYLLEEFRTKGAPIDPTFFLSRTVS 179

Query: 180 NVVCSLLFGLRFSYEDKEFQAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLH 239
NV+ S++FG RF YEDK+F +++R + + +S+ Q Y+M+S ++ LPG H ++ +
Sbjct: 180 NVISSVVFSGRFDYEDKQFLSLLRMINESFIEMSTPWAQLYDMYSGVMQYLPGRHNRIYY 239

Query: 240 HVSTLAAFTVRVQVQHQGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYL 299
+ L F +V+ ++ +LD P RD +D FL+KM Q++ NP TEF KN+++T + L
Sbjct: 240 LIEELKDFIAARVKVNEASLDQPQP-RDFIDCFLIKMHQDKNNPHTEFNKLNVLTTNL 298

Query: 300 LFAGMTVSTTVGYTLLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTLPYTDAVLHE 359
FAGT TVS+T+ Y LL+MK+P VQ + EE+N+ +G + PS+ DR ++P+TDAV+HE
Sbjct: 299 FFAGTETVSSTLRYGFLIMKHPEVQTKIYEEINQVIGPHRIPSVDORVKMPFTDAVIHE 358

Query: 360 AQRLLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEEFNPDRFLD 419
QRL +VPMG+P ++R T FRGY LP+GT+VFPLLS+L DP F HP++F P FLD
Sbjct: 359 IQRLLTDIVPMGVPHNVIRDTHFRGYLLPKGTDVFPLLGSVLKDPKYFCHPDDFYQHFLLD 418

Query: 420 ADGRFRKHEAFLPFSLGKRVCLGEGLAKAELFLFTTILQAFSLSPCPDPTLSLKPTVS 479
GRF+K+EAF+PFS GKR+CLGE +A+ ELFL+FT+ILQ FSL PP + + P +S
Sbjct: 419 EQGRFKKNEAFVPFSSGKRICLGEAMARMELFLYFTSILQNFSLHPLVPPVNIDITPKIS 478

Query: 480 GLFNIPPAFQL 490
G NIPP ++L
Sbjct: 479 GFGNIPPITYEL 489 (SEQ ID NO:4)

Hmmer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00067	Cytochrome P450	594.4	6.9e-175	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00067	1/1	33	493	..	1 497 []	594.4	6.9e-175

FIGURE 2C

1 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
51 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1901 NNNNNNNNNN NNNNNNNNNN NNNNNNTGACA GGGGCCATGA TGGAGACACC
1951 TTGGATCGAA GAGGTCACAG CACCTCTCTC TTCTCTCTCT CCTACCCCA
2001 GCTGAGTAAG AAGTACGGAC CGGTGTTTAC CATCTACCTG GGACCTCTGGC
2051 GGCCTGTGGT GGTCTGGTT GGGCAGGAGG CTGTGCGGGA GGCCCTGGGA
2101 GGTGAGGCTG AGGAGTTCAG CGGCCGGGGA ACCGTAGCGA TGCTGGAAGG
2151 GACTTTTGAT GGCCATGGTA AGTCAAGGGC TGCTAGGCC TCCGCTCACA
2201 GCCTGCCACC ACTTACTGGT GTGTGACCTT TGCACATGGC TTAGTCCCTC
2251 TGTTGCCTCA TCTGTCAAAT GGAGTGATAA CAGTGCCCAT CAGCCGGGTG
2301 CAGTGGCTAG TGCCTGAAAT CCCAACACTT TGGGAGGCGG AGGTGGGTGG
2351 ATCACTTGAG GTGAGGAGTT CGAGACCAGC CTGGCCAACA TGGTGAAACC
2401 CTGTCTCTAC TAAAAATATA AAAATTAGCT GGGCATGGTG GTGCGTACCT
2451 GTAATCCAG ATACTTGGGA GGTTGAGGCA GGAGAATCGC TTGAACCCGG
2501 GAGGCAGATG TTGCAAGTAA CCAAGACTGT GCCACTGCAC TCCAGTCTGG
2551 GCAACAGAGT GAGCCTCCAT CTCAAACAAA CAAACAAAAA GCAGTGCCCA
2601 TCATGTAGGA TTGAGTGATT GAGTGAGGAC TGAGCCTTGT GCAAAGTGAG
2651 CACTCACTAA TCACCAGGTT GTAGTATCAG TGATAACCAT CAATGATCCA
2701 GGTAAAGCCC TAGGGGTTCA GAAAGATGCC GGAGCGCTTT CAAGGTGCTG
2751 GGGATTGGTG GGCAAGCCCT CGAATAATAG AAACAGTTCT CTGTATTACA
2801 ACAGAAAGCA GGAGGCCCAT GCTGGGTGCT GCCAGGAACT CAGTAGTAAC
2851 TAAGACAGCA CCGGTGCTGC TTCCCAGCG CACCTAGGCC AGTGGGGA
2901 CAGACTCACC ACACAGTCCC AGCCAGAGT GGTGAGGGCC AAGATGGGGA
2951 AGCAGGGGGA GAAAGGTCAG GGTGGGATGG GGAGGGGTCA GGGCAAGAGG
3001 GGTGAGGGCC AGGCTGAGGG AAGCCCTGGG ACTGTAGGAA TTTAGAGGAG
3051 GTACCTGACC CGGCATGTTT GGTGAGGGAG ATTGAGGAAG TCTTCTGGGA
3101 AGAGAGGCTG TCGGAGCTGA GACTCATAAG ATGAGTGGGG AGGGTGTTC
3151 AGGCAGAAAG ACCAGCACCT ACAAAGCAT GACTTTGAGA GAAGCATTCA
3201 TCCATTCAAC TGATGAATTT TCAGACTGGG CACGCTGGCT CATGCCTGTA
3251 ATCCCAGCAC TTTGGAAGGC TGAATGGGGA GGATGACTTG AGCCTAGGCA
3301 TTTGTGACAA CCCTGGGCAA CATGGTGAGA CCCTGCCTCC ACAAACAAAA
3351 CAAACAAACA AAAAATCATT ATACCTGGTA CCATGGGTAC CAGGTACATA
3401 GAAATGACTC AGGCAGATAT GGTGTCCTCT CCTACTGTGG GAGAGGCGGG
3451 CTTATACTGC AGTAAGACAA TAGAGGGAGG GAATATAATC CTAATATGAG
3501 AGGTACAGAT TTGAGAGCAA ACACAGGGCA CAGGCATATG TACGAGGGTA
3551 AAGAGGGAAT CAGGGAAGGC TTCTCAGAGA AGGTGACATT TAAGCCGGGA
3601 CATGAAGGAT GAACGAGTTA GTTCACCAAG GATGGGATGG AAAGGGGTGA
3651 GAGTGATGGA GGCAGAGGGA ACTGCAGGAT CATAGGCCTA GACAGGGGAT
3701 CCTGACGCCC TTGAGGAAGT GAGAGAAAGC CAGCGCAGTC GTAGTGGGTT
3751 AAGTAACAAA GCTGAGAAGC CAGGGAATTC CCTGGTCATG CAGGCTCTGT
3801 GAGTCACGTC AGAGTGTTTG GGCTTTTGT TTTCTGGGA GCAGTCGATT
3851 TTAAGCAGGG AACAGCTGTA TTCAGAGTTG GGAAGATCCT GTGGTTGCTG

FIGURE 3A

```

3901 CCTGAAGGGG ATGAACTGG AGGCTAGGAG CCCAGGGTGA TAGGGAGGAT
3951 CCAGGGTGAT GGGGAGGCTG GGAGGTCCGC GGTGATGGAC CAGGGCTGGG
4001 GCCAGGGGAT GGGGAGGAAG GAGTAATTGG GAGAGGCCTG GGGCTCTGGC
4051 CGAGGAATGG ATGGTGGGCT GAAACAGGGA GAGGAGAGAT GCTTAGGCCA
4101 CTTTGGAAACA CAGTAGGGCA AGGACAGGAG ACACCCAAGG GGAAGTGCCC
4151 AAGAGACCAC GACAGGCTGG CATTGGACAG GGAAGGTCTG TCTGGAGCAG
4201 GTGTCTTGGA TAAGGAGGGA AAATGGTGCA GTTCCATCCT CCTCCCTCTC
4251 TGTTCAACCT CTAACACTACA TGGGGCACAG GACCCAGTGG GACTCCATAA
4301 ATGATGGGAT GGGTGGATGG AAGGAAGGAA GGAGGAAACA ACTCTTCATT
4351 CATCTGGTTT ATTTACAGAA CAGGCCAGGT GCGGTGCTCA CGCTTGCCAT
4401 TCTAGCACTT TGGGAGGCTG AGGTGGGTGG ATTACCTCAG GTCAGGAGTT
4451 CAAGACCAGC CTAGACAACG TAGAGAAACC CCATCTCTAC TGAAGATATA
4501 AAATTAGCTG GGCCTAGTGG CATATGCCTG TAATCCCAGC TAGTCGGGAA
4551 GCTGAGGCAG GAAATCGCT TGAACCCGAG AGGCAGAGGT TGGCGTGAGC
4601 TGAGATCGTG CATTGCACT CCAGCTGGG TGACAAGCA AGACCTCGTC
4651 TCAATAATAA TAATAATTAC AAAACAGAAG GAGCCTGGGT CATCCCAGCT
4701 ACCTACTTTT CAGGAGAATG TACTCCCTTA CCCAAGGGCA AAGGATGGGA
4751 GAACCACTTT GATTATGCAT TTATTGAGCA CTTACTGAGT CCTCATCCCT
4801 GGGCTAGGCT GGAATGGACT CAGATGGAGC CTGAAGAGTC CCCCTCAGGG
4851 AACCTCACTA GAAAGAAGGA GGAATCGGCC GGGCGCGGTG GCTCACGCCCT
4901 GTAACTCCCA CACTTTGGGA GGCTGAGGTG GGTGGATCAC AAGGTCAGGA
4951 GATCGAGACC ATCCTGGCTA ACACAGTGAA ACCCATCTC TACTAAAAAT
5001 ACAAAAAATG AGCCAGGCAT GGTGGCGGGC GCCTGTAGTC CCAGCTACTC
5051 AGGAGGCTGA GGCAGGAGAA TTGCTTGAAC CCGGGAGGCA GAGGTTGCAG
5101 TGAGACGAGA TCACGCCACT GCACTCCAGC CTGGGCAACA GAGCGAGATT
5151 CCGTCTCAAA AAAAAAAGA AAGAAAGGAA GAAGGGGAA TGGGGGAGAG
5201 GAGCCGGTCC CTTTTTGAGT CTAGCCTTCT GCGCAGGGGT TTTCTTCTCC
5251 AACGGGGAGC GGTGGAGGCA GCTGAGGAAG TTTACCATGC TTGCTCTCGC
5301 GGACCTGGGC ATGGGGAAGC GAGAAGGCCA GGAGCTGATC CAGGCGGAGG
5351 CCGGTGTCT GTGGAGACA TTCCAGGGGA CAGAAGGTCA GCATGGCGGG
5401 GTCACCCAG GGTCTCCAGC CGAGTGAAAG GGAAACTCT CTTACTGTGG
5451 CTGGGGGTGG CCCCAACCCA GGTCTGGAA TGGGCAGGAG GGAAGCCTT
5501 GAACCTAGG GCTGGCCTGG GGGTTCTGTT CACTGCCACC TTCTGTCTCT
5551 GTCCCACTGT CTCTCCGAGG CTGTCTGATC ATCTCTCTGT GTGTCTCTGG
5601 TGCTATCATC CATTCTTCC TGGGTCTCCA TCTCTCTCT TGTCTCTTTT
5651 CTTTCTCTCT CTTTCTCTCT ATTTTGTGG CCCTCAGTCT ATCTCTGTTT
5701 CTGTCTCCCT GTCTGTGTGA TGGTCACTCT GTTCTTTTCT CCCTGTCTGT
5751 TTCTCTGTCC CTATCTGTCT GTATCCTTCT TTGCTGTGTT AGCTCTCTCC
5801 CTGCGCTGTC CATCCATCTT TCCCTGCCTC CCTGTCTCTC TCTGTTGGG
5851 TTCAGCCCCA ACCTGCTCCC CTCTGCCTGG CTCCATCACA GCCTACCTCC
5901 CTGCCCCCAT TCCCCCAGG ACGCCCATTC GATCCCTCCC TGCTGTGGC
5951 CCAGGCCACC TCCAACGTAG TCTGCTCCCT CCTCTTTGGC CTCCGTTCT
6001 CCTATGAGGA TAAGGAGTTC CAGGCCGTGG TCCGGGCAGC TGGTGGTACC
6051 CTGCTGGGAG TCAGCTCCCA GGGGGGTGAG GTGAGTGGGT GGGACCCCTC
6101 TCCAACCTACC TTCCCTGAAG GTTCTGCGCA AGGTCCCAGT AGAACTAGCT
6151 GCCCTTCTCC CCACAGACCT ACGAGATGTT CTCCTGGTTC CTGCGGCCCC
6201 TCCAGGGCCC CCACAAGCAG CTCCTCCACC ACGTCAGCAC CTTGGCTGCC
6251 TTCACAGTCC GGCAGGTGCA GCAGCACCAG GGAACCTGG ATGCTTCGGG
6301 CCCCACACGT GACCTTGTCG ATGCTTCTCT GCTGAAGATG GCACAGGTGT
6351 GGGAAAGGGT CAGGGACCCC CTCTCTGAAT GGGCGTGGTG ACCTGGCAGG
6401 TCCCAGCCAG GTGTCCCTGG GGACCTCAAT TGGGTTCCCT TCTCTTCTC
6451 TCTCTGCATG TCTCTGTGAG TATGAGTGTCT TCTGTGCATG TGTGTGCATC
6501 CCTTCTCTGC ACATCTGTGC TGGCCCTTTC AGGGCGTTGC TCTCACTGCC
6551 TCTCCCGCCC CCGACCTGGG CATTTGTGCC GGGCTGTCTG TCTCTCCAGC
6601 ATCTCTCCTC TTTCTCCCTC CCACCTCGGC CCTTGTGTTT AGGCCCCATG
6651 CCCAGGGTCC TACACCAGCA ATCCCCAGGA TCACCTCATC CCATCCCCTG
6701 CAGCCTCCCC AGACTTTTAT GTAAATTCAC AATTTTATGT GAATTATGGT
6751 CATTTATTAG GAAGCCTTGC AATATCAAGT TATGTTAATA AAGTCCACTT
6801 TATTAATTAT ATAAGAACA TATTTCTTTT CCTTTTCTTT TTCTTTTCTT
6851 TTTAAAGAGA CAGGATCTCT TTCTGTTGCC CAGGCTAGAG TACAGTTGCA
6901 AAATCATAGC TCACTGCAAC CTTGAACCTC TGGGCTCAAG CAATCCTCCT
6951 GCCTCGGGCT CCTGAGTAGC TGGGACAACA GGTGTGCACC ACCACACCTG
7001 GCTAAATTTT TTTTTTTTCT TTGTAGAGAT AGACTCTCAC TATGTTACCC
7051 AGGCTGGTCT TGAATTCCTG GGCTCATGTA ATCCTCTGTC TGCTTGAAC
7101 TCCCAAAGTG CTGGGACTAT AGGCATAAGA CATCATGCCC GGTGGGCAC
7151 AGTGCTCATC GCCTGTAATC TCAGGACTTT GGGAGGCCGA GACGGGCGGA
7201 TCACCTGAGG TCGGGAGTTC GAGACCAGCC TGACCAACAT GGAGAAACCC
7251 CATCTCTACT AAAAAAAAAA ATACAAATTT AGCCGGACGT GGTGGCACAT
7301 GCCTGTAATC CCAGCTACTA GGGAGGCTGA GGCAGGAGAA TCGCTTGAAC
7351 CCGGAGGCT TAGGTTGCGG TGAGCTGAGA TTGCACCATT GCACTCCAGC
7401 CTGGGCAACA AGAGCGAAAT TCCATCTCAA AAAAAAAAAA AAAGAAAAAA
7451 AGAAAAAAGA CACCATGCCC TATAAGTAAA CTAGAATTAA GGTGACTCCT
7501 AAGGAAATAA ATAGTTTTTA ACTGTACGAA CTTTTGGAAG AATGGGGCCA
7551 ATTCTTTAAT TAAATGCAGC CTCCCTGTTT GTGGAGAAAG AAAAAATTTT
7601 CTTAACCCAT TTGCCCCATT TCTTTTCTCT TTTATTGAAT ATTTTTTAGT
7651 TTTAACTATA GTAAAATACA CATAACGTTT ACCATCTTAA CCATTTTTAG
7701 GTATACAGTA CAGTAGTGT CAGTACATTC ATACTGTTAT GCAATCAGTC
7751 TCCAGAACTC TTCATGTTGC AAAGCTGAAA CTCTATACCC ATTAACAAC

```

FIGURE 3B

7801	TGCTGTTC	TCCCTCTCC	AACCCCTGGC	AATCACCTTT	TTTTTTTGA
7851	GACGAAGTCT	CACTCTGTCA	CCCAGGCTAG	AGTGCAGTGG	CTCGATCTCG
7901	GCTCACTGCA	AGCTCCGCT	CCCGGTTCA	TGCCATTCTC	CTGCCTCAGT
7951	CTCCAAGCA	GCTGGGACTA	CAGGTGCCCG	TCACCACGCC	TGGCTAATTT
8001	TTTGTATTTT	TAGTAGAGAT	GGAGTTTCAT	CGTGTTAGCC	AGGCTGATCT
8051	CAAACCTCTG	GCCTCAAGTG	ATCCACCCGC	CTCGGCTCC	CAAAGTGCTG
8101	GGACTACAGG	CGTGAGCCAC	TGTGCCTGGC	CAGGAAGTAG	ACTCTTGATA
8151	TTAGTTCTCT	CTGGTTGAAA	TGTTTTTAAA	AATGAAAGAG	AATGACTAAT
8201	AACAAAAACA	CAGAAAGTTA	TAAGGATTGA	TGAAGATGTG	GAGACTTTGA
8251	AACCCATGTA	TACCATTGGT	GGGAATGTGA	AACGACGCAG	CCCTGTGGAA
8301	AATGGTACAG	CAGTTACCTG	AGGTCAGGAG	TTTGAAACCA	ACCTGGCCAA
8351	CATGCAGAAA	CCCCGTCTCC	ATTAAATGTA	CAAAAATTAG	CCAGGNNNNN
8401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN

FIGURE 3C

11701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12601 NNNNNNNNNN NNNNNNNNNN NNNNNNNCTCT TTCTTCTCTT CTCTCTCTCT
12651 CCCTTCTCC CTCTCTCCCT TCCTGTCTTC CTCTCTTCTT CTCTTCTTCT
12701 CTTGACAGG TCTCTCTTTG TCTCCCAGGC TGGGGTGCAG TGGTACAAGC
12751 ATAGCTCACA GCAGCCTTGA ACTCCTAGGC TCAAGTGATC CTCCACGTC
12801 AGCCTCCTGA GCAGCTGGGA CAACGGGCTC ATACCACCAT GCCTGGCTAA
12851 TTTTATAATT TTTCTAGAG ACAAGGCTCT GTTATATTGC CCAGGCTGGT
12901 CTCAAACCTC TGGGCTCAAA TGCTTCTCTC ACCTCAGCCT CCCACGTGGC
12951 TGGGATTACA GGCATGAGCC ACTGCACGCC ACTCAACACT CCACAAATGT
13001 TGATGCCATT ATGTTTTGTG AACTAGTGTG CCTGGCACCC GAGACTTGTA
13051 CTCCACACTC GAGGACCAAA TAGACTGGGG TGGGAAGGGG TTTATAGTTT
13101 CATTATTATT TCCCCTCAGG GCACGGAGGT CTTCCTCTCT CTCTGGCTCA
13151 TCCTGCATGA CCCCACATC TTCAAGCACC CAGAAGAGTT CAACCCAGAC
13201 CGTTTCTCTG ATGCAGATGG ACGGTTACAG AAGCATGAGG CGTTCTCTCC
13251 CTCTCTCTTA GGTATCTGCT GCAGCCCTGG GTATCACAAG CAGGTGCTGG
13301 CGAACTCCAG GCATCTGTGC CAGCTGGGGG CACCTTCTCT CACCCTGGGC
13351 TTAAGTGTG CTCTCTCACC TGCTGTCTCC CCCGTGGGCC TGGGTGTGAG
13401 GAATACTGAC TCAGCCCTCT CTCTCTCTCT CTCTCACC GGAAGCGTG
13451 TCTGCTTGG AGAGGGCCTG GCAAAGCGG AGCTCTCTCT CTCTTCAACC
13501 ACCATCTCTA AAGCCTTCTC CCTGGAGAGC CCGTGCCCGC CGGTACACCC
13551 TGAGCCTCAA GCCCACCCTC AGTGGCCTTT TCAACATTCT CCCAGCCTT
13601 CAGCTGCAAG TCCGTCCAC TGACCTTAC TCCACCACGC AGACCAGATG
13651 AAGGAAGGCA ACTTGGAGT GGTGGGTGCC CAGGACGGTG CCTCCAGCT
13701 CAACAGTGGG CATGGACAGG GTTAATGTCT CCAGAGTGTA CACTGCAGGC
13751 AGCCACATTT ACACGCCTGC AGTTGTTTTT CGGAGTCTGT CCCACGGCCC
13801 ACACGCTCAC TTGACTCATG CTGCTAAGT GCACAACCGC ACACCCATAC
13851 ACAACTACAA GGGCCACAAA GCAACTGCTG GGTTAGCTTT CCACAGACAT
13901 AAATATAGTC CATCTGCAAT CACAAGCACA TAGCCAGGTA ACCCACCAC
13951 TCCCCTGGAT CTGCAGCCCA CACGTGGGAG TCTGGCTGTC ACCTTCACAA
14001 GCCACAGAAA CGGCCACACA TGTTACAGC TCACACGCCC TCTCCATTCA
14051 TCGAATCTCT CAGTGTCCCT GTCCCTGGTG CCTGGCACAG GGAACAGCAT
14101 GCCCCTCCG GGGTCATGCC ACCCAGAGAC TGTCTGCTGCT TATGGCCCCA
14151 ACTCATGTCT CCTCTCTTGG CTACACCACT CTCCAGCCT GTGACCACCG
14201 ATGTCCACAC ACCCCCAACC ACTTGTCCAC ACAGTACCC ACCTACGACA
14251 TCGTCTGGC TCCCAGAGT ATCTTCCAC TGAGACACGC CGCCCCACA
14301 GAGGCACAGT CCCCAGCCAC CTCTGCAACT GCAGCCCTCA GTCAACCTT
14351 TTTAAGCACC CTGATTCTAC CAAATGCAAA CACATCTGGG TCTGCGATTA
14401 TGCACAGAGA CTTTGGACAT ACGAGGACCC TCAGACCGGA GGAACACCTG
14451 CCAACCCCA ACACGTGCTT ATGTAACCAC GTGGAAGCG GCCCCTGCTG
14501 CCCCTCCACA CACACATACA CACTCACTGA TCTACAGCCC CTGTTCCGCG
14551 TCAGAGTCCC CACTAGACCC AGTGGAAGGG GTTAGAGACC AAGTAGGGGC
14601 CAGTTTCAA TTAACCTGT CAGGGAGTGA GCCGATCTG ACGTTCCTTG
14651 TGACTTAAGG GTCCGGCTTG GGAATTAAG TTTGTTCTG GCCTTAGCC
14701 TACTGCGTGT GTGACCCGTG TCAGTCACTG TGAGTAAGGG GTGGGACAG
14751 GGGAGTCCAC CCTTCCCTG AGGCTGGGCG GGAGTGAAA AACATGGCCA
14801 CCGCCACCC TGGCTGTTGA CATCAGGACC AGATGTGGAG CTGGGAGGAG
14851 GGGCAGGGCT GGTGACGCC TGGCCTCAT TTCCAAAAG GGCACAGGTG
14901 TCCGGCGGTG GGAAGTGGG AAGGAGGGG TAACCCAAGC TGGACTGTGG
14951 ACCTTGGGG CTCTCTCAGC CAGGAGAGC CTGAAGCCAA CTAGATCCAG
15001 ACCCTAGAGA CTCTTCAAAC TTGAGTACAG GAACTAGCTT GCAACACAGA
15051 CTCTAAGCCC ACTCCCATTT CTCTCACCT TTTTCTCTTG CCTCCCTTC
15101 ACAAGGAAAC CAGAGCATT TGTAATTTT CTCTCTTTT TTTTCTTTT
15151 TTTTCTTTGA GACGGAGTCT CACTCTGTCA CCCAGGCTGG AGTGACGTGG
15201 TGTGATCTTG GCTCACTGCA GCCTCCGCT CCGGTTCAA GCCATTCTCC
15251 TGCTCAGCC TCCCAAGTAG CTGGGATTAC AGGTGTGTGC CACCACGCCC
15301 AGCTAATTT TGTATTTTA GTAGAGATGG GGTTCACCA TGTGGCCAG
15351 GCTGGTCTCG AACTCTGAC CTCAGATGAT CTGCCAGTCT CGGCCTCCCA
15401 AAATGCTGGG ATTACAGGCG TGAGTCGCTA CTAGATAAAT TTCTTATCTA
15451 GCAAAGAAGT TTGCAACAT ACGCAAAAGT AGAAAGATAC AATGAGCCCC
15501 CAGGTGCCCA TCACCCAGCC TCATTTCAAT AGTCATCAAC TTTCTGCAGC
15551 TTTTACTTCA TCTATATCCT TTTCTGCCTC TTTTCTTTT TTTTATTTT

FIGURE 3D

```

15601 GAGATAGGGT TTTGCTTTGT TGCCCAAGCT GGGGTGCAGT AGCATGATCT
15651 CATAGTTCAC TGGGGCTTCA GACTCCTAGG CTCAAGTGAT CCTCCCGCCT
15701 CCGCCTCCAA GCAGCTGGGA CTACAGATGC GTGCCACCAC ACCCAGCTAA
15751 ATTTCTTATT TTTATTTTCT ATAGAGAAAG TCTCACTATA CAGCCCTGTG
15801 CTGGTCTCAA ATTCCAGGCC TCAAGAGTTT CCATCCCAGC CTCCCAAAGT
15851 GGTGGGATTA TAGGCGTGAG TCACTGCACC CTGCCCTAAT ATTTTATTTT
15901 TATCTATTGC TTTTATTTTA CTTATTTATC TTTTATTTT GAGACAGAGT
15951 CTCACTCTGT GGCCCATGCT GGAGTGCACT GGCATCATCT CGGCTCACTG
16001 TAACCTCCGC CTCTTAGGTT CAAGCAGTTC TCCTGCCTTG ACCTCCCGAG
16051 TAGCTGGAAT TACAGGTGCC TGCCACCAAG CCTGGCTAAT TTTTATTTT
16101 GTAGTAGAGA TGGGGTTTTG CCATGTTGAC CAGGCTGGTC TCGAACTCCT
16151 GACCTCAGGT GATCTGCCCA CCTTGGCCTC CCAAAGTGCT GAGATTACTG
16201 GTATGAGCCA CCGTGCCTGG CCACCTATTG CTTTTTAAAG ATTATTTTTT
16251 TATTATTATT ATTTTTTTAT TTGCAGATGG AGTTTTGCTA TTGTTGCCCA
16301 GGCTGGAGTA CAATGGCGTG ATCTCAGCTC ACCGCAACCT CCGCCTCCCA
16351 GGTTCAGCG ATTCTCCTGC CTCAGCATCC CTAGTAGCTG GGATTACAGG
16401 CATGCACCAC CATGTCAGC TAATTTTGT TTTTATAGTAG AGACGAGGTT
16451 TCTCCAGGTT GGTCAAGGCT GTCTCAAAC CCAAACCTCA GGTGATCCGC
16501 CCATCTCGGC CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACCGCGCC
16551 TGGCCTTAAA GATTATTTTA AGGCAAATTA CAGAAAGCAA TTTAATGCAC
16601 ATTTCTGAGA GTTAAAGATA TTTTGGCCT TGACATTTTA TGAGGACAGT
16651 TTTCAAACAT GCAGCAAAGT TGAGGGAATT GTACAAGGAA CACCTTGTGC
16701 ACTCCTGCCT CAGTCTCCCA AGCAGCTGGG ACTACAGGTG CCCGTCACCA
16751 CGCCTGGCTA ATTTTGTGTA TTTTATAGTAG AGATGGAGTT TCATCGTGT
16801 AGCCAGGCTG ATCTCAAAC CTGGCCTCA AGTGATCCAC CCGCCTCGGC
16851 CTCCCAAGTG CTGGGACTAC AGGCGTGAGC CACTGTGCCT GGCCAGGTAA
16901 GTAGACTCTT GATATTAGTT CTCTCTGGTT GAAATGTTTT TAAAAATGAA
16951 AGAGAATTGAC TAATAACAAA AACACAGAAA GTTATAAGGA TTGATGAAGA
17001 TGTGGAGACT TTGAAACCCA TGTATACCAT TGGTGGGAAT GTGAAACGAC
17051 GCAGCCTGTT GGAAATGGT ACAGCAGTTA CCTGAGGTCA GGAGTTTGAA
17101 ACCAACCTGG CCAACATGCA GAAACCCCGT CTCCATTAAA TGTACAAAAA
17151 TTAGCCAGGC ATGGTGGTGC GCACCTGTAA TCCCAGCTAC TCGGGAGGCT
17201 GAGGCAGGAG AATTGCTTGA ACCCAGGAGG CGGAGGTGTC AGTGAGCCGA
17251 GATCGTGCCA CTGCACTCAG CCTGGGCAAC AAAGCAAGAC TCTGTCTCAA
17301 AAAAAAAAG TCTACTTCCC AACCTTCCC AAAATTTATC TAAACCCCGT
17351 GACAAAACCT TAACTTGTGT TTCCGACCCC AGGCTTGGCT GTTCTGGACA
17401 TTTACTTCCC AAAGGCTGTG TTCTCTCAGC CCCTCTGCCT GGTTTCTTTC
17451 AGGAGGAACA AAACCCAGGC ACAGAATTCA CCAACAAGAA CATGCTGATG
17501 ACAGTCATTT ATTTGCTGTT TGCTGGGACG ATGACGGTCA GCACCACGGT
17551 CGGCTATACC CTCCTGCTCC TGATGAAATA CCCTCATGTC CAAAGTAAGA
17601 GCCTTTTCCA CTTGCCAGGC CTTGGGAACA GAAGTCAGGG TTCTAGGCTG
17651 AGCAAGGTGG CTCACGCCTA TAATCCAGC ACTTTGGGAG GCTGAGGCGG
17701 GCTGATCACT TGAGAATAGG AGTTTAAGAC CAGCCGCCCA ACACAGTGAA
17751 AC (SEQ ID NO:3)

```

FEATURES:

Start:	1999	6349
Exon:	1999	2167
Intron:	2168	5236
Exon:	5237	5386
Intron:	5387	5919
Exon:	5920	6080
Intron:	6081	6166
Exon:	6167	6349
Intron:	6354	13542
Exon:	6354	6636
Intron:	6637	13087
Exon:	13088	13261
Intron:	13262	13441
Exon:	13442	13542
Intron:	13547	13648
Exon:	13547	13648
Stop:	13649	

SNPs:

DNA Position	Major	Minor
2226	G	C A
2226	C	A G T
3081	A	T G C
3788	A	T
3979	T	C G
5056	T	C G
5213	T	G A C
5508	A	C T G
5857	C	G A T

FIGURE 3E

6385	C	T G
6813	C	T A
7853	C	T
12973	G	C T
12973	A	G T
13012	G	A C T
13072	A	T C G
13370	C	G
13682	C	T A
14631	C	T G

Context:

DNA
Position

2226	TGACAGGGGCCATGATGGAGACACCTTGGATCGAAGAGGTCACAGCACCTCCTCTTTCT TCCTCCCTACCCCAGCTGAGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACC CTGGCGGCCTGTGGTGGTCCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTCA GGCTGAGGAGTTAGCGGCCGGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCA TGGTAAGTCAAGGGCTGCTAGGCCCTCCGCTCACAGCCTGCCACCACTTACTGGTGTGTG [G,C,A] CCTTTGCACATGGCTTAGTCCCTCTGTTGCCTCATCTGTCAAATGGAGTGATAACAGTGC CCATCAGCCGGGTGCACTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGG GTGGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCT CTACTAAAAATATAAAATTAGCTGGGCATGGTGGTGCCTACCTGTAATCCCAGATACTT GGGAGGTTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGA
2226	TGACAGGGGCCATGATGGAGACACCTTGGATCGAAGAGGTCACAGCACCTCCTCTTTCT TCCTCCCTACCCCAGCTGAGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACC CTGGCGGCCTGTGGTGGTCCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTCA GGCTGAGGAGTTAGCGGCCGGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCA TGGTAAGTCAAGGGCTGCTAGGCCCTCCGCTCACAGCCTGCCACCACTTACTGGTGTGTG [C,A,Gst]

FIGURE 3F

CCTTTGCACATGGCTTAGTCCCTCTGTTGCCTCATCTGTCAAATGGAGTGATAACAGTGC
CCATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGG
GTGGATCACTTGAGGTCAGGAGTTTCGAGACCAGCCTGGCCAACATGGTGAACCTGTCT
CTACTAAAAATATAAAATTAGCTGGGCATGGTGGTGCCTGTAATCCCAGATACTT
GGGAGGTTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGA

3081 AAACAGTTCTCTGTATTACAACAGAAAGCAGGAGGCCATGCTGGGTGCTGCCAGGAAC
CAGTAGTAACCTAAGACAGCACCGGTGCTGCTTCCAGCGCACCTAGGCCAGTGGGAAA
CAGACTCACCACACAGTCCAGCCAGAGTGGTCAGGGCCAAGATGGGGAAGCACGGGA
GAAAGGTGAGGTTGGGATGGGAGGGGTGAGGGCAAGAGGGGTGAGGGCCAGGCTGAGGG
AAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGACCCGGCATGTTTGGTGAAGGAG
[A, T, G, C]
TTCAGGAAGTCTTCTGGAAGAGAGGCTGTGAGGCTGAGACTCATAAGATGAGTGGGA
GGGTGTTCCAGGCAGAAAGACAGCACCTACAAAAGCATGACTTTGAGAGAAGCATTAT
CCATTCAACTGATGAATTTTCAGACTGGGCACGCTGGCTCATGCTGTAATCCCAGCACT
TTGGAAGGCTGAATGGGAGGATGACTTGAGCCTAGGCATTTGTGACAAGCCTGGGCAAC
ATGGTGAGACCTGCTCCACAAAACAAACAAACAAACAAACAAATCATTATACCTGGTAC

3788 ATCTTAAATGAGAGGTACAGATTTGAGAGCAAACACAGGGCACAGGCATATGTACGAGG
GTAAGAGGGAATCAGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCCGGGACATGAAG
GATGAACGAGTTAGTTACCAAGGATGGGATGGAAGGGGTGAGAGTGATGGAGGCAGAG
GGAACCTGAGGATCATAGGCCAGACAGGGGATCCTGACGCCCTTGAGGAAGTGAAGAA
GACCAAGCGAGTCGTAGTGGTTAAGTAACAAAGCTGAGAAGCCAGGGAATCCCTGGTC
[A, T]
TGCAGGGCTGTGAGTCACGTGAGTGTGTTGGGCTTTTGTCTTCTGGGAGCAGTCGA
TTTTAAGCAGGGAACAGCTGATTGAGTGGGAAGATCCTGTGTTGCTGCTGAAGG
GGATGAACTGGAGGCTAGGAGCCAGGGTGATAGGAGGATCCAGGGTGATGGGAGGC
TGGGAGGTCCCGGTGATGGACAGGGCTGGGGCCAGGGGATGGGAGGAAGGAGTAAT
GGGAGAGGCTGGGGCTCTGGCCGAGGAATGGATGGTGGCTGAAACAGGAGAGGAGAG

3979 ATCATAGGCCTAGACAGGGGATCCTGACGCCCTTGAGGAAGTGAAGAAAGACAGCGCAG
TCGTAGTGGTTAAGTAACAAAGCTGAGAAGCCAGGGAATCCCTGGTATGACGGGCT
GTGAGTCACGTGAGTGTGTTGGGCTTTTGTCTTCTGGGAGCAGTCGATTTTAAGCAG
GGAACAGCTGATTGAGAGTTGGGAAGATCCTGTGTTGCTGCTGAAGGGGATGAACT
GGAGGCTAGGAGCCAGGGTGATAGGAGGATCCAGGGTGATGGGAGGCTGGGAGGTCC
[T, C, G]
CGGTGATGGACAGGGCTGGGGCCAGGGGATGGGGAGGAAGGAGTAATGGGAGAGGCT
GGGGCTCTGGCCGAGGAATGGATGGTGGGCTGAAACAGGGAGAGGAGATGCTTAGGCC
ACTTTGGAACACAGTAGGGCAAGGACAGGAGACCCCAAGGGGAAGTGCCTAAGAGACCA
CGACAGGCTGGCATTGGACAGGGAAGGTCTGTCTGGAGCAGGTGTCTTGATAGGGAGG
AAATGGTGCAGTTCATCCTCCTCCTCTCTGTCAACCTCTAACTACATGGGGCACA

5056 AGTTTGATTATGCATTTATTGAGCACTACTGAGTCTCATCCCTGGGCTAGGCTGGAAT
GGAATCAGATGGAGCCTGAAGAGTCCCTCAGGGAACCTCACTAGAAAGAAGGAGGAAT
CGGCCGGGCGCGGTGGCTCACGCTGTAATCCCAACACTTTGGGAGGCTGAGGTGGGTGG
ATCACAAGGTGAGGAGATCGAGACCATCCTGGCTAACACAGTGAACCCCATCTCTACTA
AAAATACAAAAAATGAGCCAGGCATGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG
[T, C, G]
CTGAGGCAGGAGAAATGCTTGAACCCGGGAGGCGAGGTTGAGTGAGACGAGATCACGC
CACTGCACTCCAGCCTGGGCAACAGAGCGAGATTCCGTCTCAAAAAAAGAAAGAAA
GGAAGAAGGGGAATGGGGAGAGGGGCGGTCCCTTTTGTGCTAGCCTTCTGCGCAG
GGGTTTTCTTCTCAACGGGGAGCGGTGGAGGCAGCTGAGGAAGTTACCATGCTTGGCTG
TGCGGGACCTGGGCATGGGGAAGCGAGAAGGCGAGGAGCTGATCAGGCGGAGGCCCGGT

5213 CTTTGGGAGGCTGAGGTGGGTGGATCACAAGGTGAGGAGATCGAGACCATCCTGGCTAAC
ACAGTGAAACCCATCTCTACTAAAAATACAAAAATGAGCCAGGCATGGTGGCGGGCGC
CTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCAGA
GGTTGCAGTGAGACGAGATCAGCCACTGCACTCCAGCCTGGGCAACAGAGCGAGATTCC
GTCTCAAAAAAAGAAAGAAAGAAAGGAGGGGAATGGGGAGAGGGGCCGTCCCT
[T, G, A, C]
TTTGAGTCTAGCCTTCTGCGCAGGGGTTTTCTTCTCAACGGGGAGCGGTGGAGGCAGCT
GAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATGGGGAAGCGAGAAGGCGAGGA
GCTGATCCAGGCGGAGGCCCGGTGTCTGGTGGAGACATTCAGGGGACAGAAGGTGAGCA
TGGCGGGGTCAACCCAGGGTCTCAGCCGAGTGAAAGGGAAAACTCTCTACTGTGGCTG
GGGTGGGCCCAACCCAGGTCTGGAATGGGAGGAGGGGAAGCCTTGAACCTGAGGGCT

5508 TCCCTTTTTGAGTCTAGCCTTCTGCGCAGGGGTTTTCTTCTCAACGGGGAGCGGTGGAG
GCAGCTGAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATGGGGAAGCGAGAAGG
CGAGGAGCTGATCCAGGCGGAGGCCCGGTGTCTGGTGGAGACATTCAGGGGACAGAAGG
TCAGCATGGCGGGGTCAACCCAGGGTCTCAGCCGAGTGAAAGGGAAAACTCTCTACTG
TGGCTGGGGGTGGCCCCAACCCAGGTCTGGAATGGGAGGAGGGGAAGCCTTGAACCT
[A, C, T, G]
GGGCTGGCCTGGGGGTTCTGTTCACTGCCACCTTCTGTCTGTCTCCCACTGTCTCCGA
GGCTGTGATGACATCTCTGTGTCTGTCTGGTGCTATCATCCATTCTTCTGGGTCTC
CATCTCTCTCTGTCTCTTTCTTCTCTCTCTCTCTCTATTTTGGGCCCTCAGT
CTATCTCTGTTTCTGTCTCCCTGTCTGTGATGGTCACTCTGTTTCTTCTCCCTGTCT
GTTTCTCTGCTCCATCTGTCTGTATCCTTCTTGGCTGTTAGCTCTCTCCCTGCGCTG

FIGURE 3G

FIGURE 3H

12973 CTGTCTTCTCTCTTTCTCTTTCTTTCTTGACAGGGTCTCTTTGTCTCCAGGCTG
GGGTGCAGTGGTACAAGCATAGCTCACAGCAGCTTGAAGTCTAGGCTCAAGTGATCCT
CCCACGTGAGCTCTGAGCAGCTGGGACAACGGGCTCATACCACCATGCCTGGCTAATT
TTTTAATTTTTCTGTAGAGACAAGGTCTTGTATATTGCCAGGCTGGTCTCAAACTCCTG
GGCTCAAATGCTTCTCTACCTCAGCTCCCACGTGGCTGGGATTACAGGCATGAGCCAC
[A, G, T]
GCACGCCACTCAACACTCCACAAATGTTGATGCCATTATGTTTTGTGAAGTAGTGCCCT
GGCAGCCGAGACTTGTACTCCACTCGAGGACCAAATAGACTGGGGTGGGAAGGGTTT
ATAGTTTCATTATTTTCCCCTCAGGGCAGGAGGTCTTCCCCTCCTTGGCTCCATCC
TGATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCAACCCAGACCGTTTCTGGATG
CAGATGGACGGTTCAGGAAGCATGAGGCGTTCCTGCCCTTCTCCTTAGGTATCTGCTGCA

13012 CTCTCTTTGTCTCCAGGCTGGGGTGCAGTGGTACAAGCATAGCTCACAGCAGCTTGAA
CTCCTAGGCTCAAGTGATCCTCCACGTGAGCCTCCTGAGCAGCTGGGACAACGGGCTCA
TACCACCATGCCTGGCTAATTTTTTAATTTTTCTGTAGAGACAAGGTCTTGTATATTGCC
CAGGCTGGTCTCAAACTCCTGGGCTCAAATGCTTCTCTACCTCAGCCTCCCACGTGGCT
GGGATTACAGGCATGAGCCACTGCACGCCACTCAACACTCCACAAATGTTGATGCCATTA
[G, A, C, T]
GTTTTGTGAAGTAGTGCCCTGGCAGCCGAGACTTGTACTCCACTCGAGGACCAAATA
GACTGGGGTGGGAAGGGTTTATAGTTTCATTATTATTTCCCCTCAGGGCAGGAGGTCT
TCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCA
ACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCCTGCCCT
TCTCCTTAGGTATCTGCTGCAGCCTGGGTATCACAAGCAGGTGCTGGCGAACTCCAGGC

13072 CTCCTAGGCTCAAGTGATCCTCCACGTGAGCCTCCTGAGCAGCTGGGACAACGGGCTCA
TACCACCATGCCTGGCTAATTTTTTAATTTTTCTGTAGAGACAAGGTCTTGTATATTGCC
CAGGCTGGTCTCAAACTCCTGGGCTCAAATGCTTCTCTACCTCAGCCTCCCACGTGGCT
GGGATTACAGGCATGAGCCACTGCACGCCACTCAACACTCCACAAATGTTGATGCCATTA
TGTTTGTGAAGTAGTGCCCTGGCAGCCGAGACTTGTACTCCACTCGAGGACCAAAT
[A, T, C, G]
GACTGGGGTGGGAAGGGTTTATAGTTTCATTATTATTTCCCCTCAGGGCAGGAGGTCT
TCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCA
ACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCCTGCCCT
TCTCCTTAGGTATCTGCTGCAGCCTGGGTATCACAAGCAGGTGCTGGCGAACTCCAGGC
ATCTGTGCCAGCTGGGGGACCCCTTCTGCACCCTGGGCTTACTGTTGGCTCCTCACCTG

13370 ATAGACTGGGGTGGGAAGGGTTTATAGTTTCATTATTATTTCCCCTCAGGGCAGGAGG
TCTTCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCACCCAGAAGAGT
TCAACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCCTGC
CCTTCTCCTTAGGTATCTGCTGCAGCCCTGGGTATCACAAGCAGGTGCTGGCGAACTCCA
GGCATCTGTGCCAGCTGGGGGACCCCTTCTGCACCCTGGGCTTACTGTTGGCTCCTCCAC
[C, G]
TGCTGTTCCCCCGTGGGCTGGGTGTGAGGAATACTGACTCAGCCCTCTCTCTCTCT
CTCCTCACCAGGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAAAGCGGAGCTCTTCT
CTTCTTACCACCATCTTACAAGCCTTCTCCCTGGAGAGCCCGTCCCGCCGGTACACCC
TGAGCCTCAAGCCCAACGTCAGTGGCCTTTTCAACATTCCCCAGCCTTCCAGCTGCAAG
TCCGTCCTCACTGACCTTCACTCCACCAGCAGACCATGAAGGAAGGCAACTTGAAGT

13682 CCGTGGGCTGGGTGTGAGGAATACTGACTCAGCCCTCTCTCTCTCTCTCTCTCACCAG
GGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAAAGCGGAGCTTCTCTCTCTTACCA
CCATCTACAAGCCTTCTCCCTGGAGAGCCCGTGGCCGCCGGTACACCTGAGCCTCAAG
CCCACCGTCAGTGGCCTTTTCAACATTCCCCAGCCTTCCAGCTGCAAGTCCGTCCACT
GACCTTCACTCCACCAGCAGACCATGAAGGAAGGCAACTTGAAGTGGTGGTGCCC
[C, T, A]
GGACGGTGCCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAATGTCTCCAGAGTGTACA
CTGCAGGCAGCCACATTTACACGCCTGCAGTTGTTTTCCGGAGTCTGTCCACGGCCAC
ACGCTCACTTGAATCATGCTGCTAAGATGCACAACCGCACACCATACAACTACAAGG
GCCACAAAGCAACTGCTGGGTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCA
CAAGCACATAGCCAGGTAACCCACCAACTCCCCTGGATCTGCAGCCACACGTGGGAGTC

FIGURE 31

14631 GCAGCCCTCAGTCACCCCTTTTAAAGCACCCCTGATTCTACCAAATGCAAACACATCTGGG
TCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCCTCAGACCGGAGGAACACCTG
CCCAACCCCAACACGTGCTTATGTAACCACGTGGAAAGCGGCCCTGCTGCCCCCTCCACA
CACACATACACACTCACTGATCTACAGCCCTGTTCCGGCGTCAGAGTCCCCACTAGACCC
AGTGAAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACCCTGTCAGGGAGTGA
[C,T,G]
CCGGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAATTAAAGTTTGTCTGG
CCTTAGCCTACTGCGTGTGTGACCCGTGTCACTCACTGTGAGTAAGGGGTGGGGACAGG
GGAGTCCACCCCTCCCCTGAGGCTGGGCGGGAGCTGAAAAACATGGCCACCGCCACCCT
GGCTGTTGACATCAGGACCAGATGTGGAGCTGGGAGGAGGGGCAGGGCTGGTGACGCCCT
GGCCTCATTTCAAAAAGGGCCAAGGTGTCCGGCGGTGGGAAGTGGCAAGGAGGGGT

Chromosomal Map Position

ePCR to dbSTS Site (bases)	Marker	Chr.	Organism
15155..15288	stSG46708	19	Homo sapiens